



IFW16

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,553D

DATE: 07/26/2004

TIME: 14:06:30

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
 Output Set: N:\CRF4\07262004\I784553D.raw

3 <110> APPLICANT: ZHOU, MING-MING
 4 AGGARWAL, ANEEL
 6 <120> TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
 8 <130> FILE REFERENCE: 2459-1-003 CIP
 10 <140> CURRENT APPLICATION NUMBER: 09/784,553D
 11 <141> CURRENT FILING DATE: 2001-02-16
 13 <150> PRIOR APPLICATION NUMBER: 09/510,314
 14 <151> PRIOR FILING DATE: 2000-02-22
 16 <160> NUMBER OF SEQ ID NOS: 63
 18 <170> SOFTWARE: PatentIn version 3.0
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 3014
 22 <212> TYPE: DNA
 23 <213> ORGANISM: Homo sapiens
 25 <400> SEQUENCE: 1
 26 gggccgcgt cgacgcggaa aagaggccgt gggggcctc ccagcgctgg cagacaccgt 60
 27 gagggctggca gcccggca cgcacaccta gtccgcagtc ccgaggaaca tgtccgcagc 120
 28 caggcgccgg agcagagtcc cgggcaggag aaccaaggga gggcgtgtgc tgtggcgccg 180
 29 gcggcagccgg cagcggagcc gctagtcccc tccctctgg gggagcagct gccggcgtg 240
 30 cccggccgc caccaccatc agcgcgcggg gcccggccag agcgagccgg gcgaggccg 300
 31 cgcttaggggg agggcggggg cggggagggg ggtggcgaa gggggcgaaa gggcgtaggg 360
 32 ggagggtctc gctctcccgaa ctaccagagc ccgaggggaga ccctggcgcc ggcggccgg 420
 33 cctgacactc ggccgcctt gccgtgtcc gggccggcat gtccgaggct ggcggggccg 480
 34 ggccggccgg ctgcggggca ggagccgggg cagggccgg gcccggggcg ctgcggccgc 540
 35 agcctgcggc gttccggcc ggcggccgc agggctcccc ctgcggccgt gccggccgggg 600
 36 gctcgccgcg ctgcggcccg gcacggcag tggctgcagc gggcacggcc gaaggacccgg 660
 37 gagggcggtgg ctgcggcccgaa atcgcgtga agaaagcgca actacgctcc gctccgggg 720
 38 ccaagaaact ggagaaactc ggagtgtact ccgcctgcaa gcccggggag tcttgtaaat 780
 39 gtaatggctg gaaaaaccct aacccctcac ccactcccccc cagagccgac ctgcagcaaa 840
 40 taattgtcag tctaacagaa tcctgtcgaa gtttagccaa tgcccttagct gctcatgttt 900
 41 cccaccttggaa gaatgtgtca gaggaagaaa tgaacagact cctggaaata gtattggatg 960
 42 tggaaatatct cttaacctgt gtccacaagg aagaagatgc agataccaaa caagtttatt 1020
 43 tctatctatt taagcttctg agaaaagtcta ttttacaaag aggaaaacct gtgggtgaag 1080
 44 gctctttggaa aaagaaaccc ccatttggaaa aaccttagcat tgaacaggggt gtgaataact 1140
 45 ttgtgcagta caaatttagt cacctgccag caaaaaggaa gcaacaata gttgagttgg 1200
 46 caaaaatgtt cctaaaccgc atcaactatt ggcatctggaa ggcaccatct caacaagac 1260
 47 tgcgatctcc caatgtatgat attctggat acaaagagaa ctacacaagg tggctgtgtt 1320
 48 actgcaacgt gccacagttc tgcgacagtc tacctcggtt cggaaaccaca caggtgtttg 1380
 49 ggagaacatt gttcgctcg gtcttcactg ttatgaggcg acaactccgt gaacaagccaa 1440
 50 gacagaaaaa agataaaactg cctcttggaaa aacgaactct aatcctact catttcccaa 1500
 51 aatttctgtc catgcttagaa gaagaagttt atagtcaaa ctctcccatc tggatcagg 1560
 52 atttctctc agccttcc agaaccagcc agctaggcat ccaaaccagg atcaatccac 1620
 53 ctccgtggc tgggacaatt tcatacaatt caacctcatc ttcccttggc cagccaaacg 1680



RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/784,553D

DATE: 07/26/2004
TIME: 14:06:30

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
Output Set: N:\CRF4\07262004\I784553D.raw

54 cagggagcag	cagtccgtcc	tgcaaaggct	cttctggact	tgaggcaaac	ccaggagaaa	1740
55 agaggaaaat	gactgattct	catgttctgg	aggaggccaa	gaaaccccga	gttatggggg	1800
56 atattccgat	ggaattaatc	aacgaggta	tgtctaccat	cacggacct	gcagcaatgc	1860
57 ttggaccaga	gaccaatttt	ctgtcagcac	actcgccag	gatgaggcg	gcaaggttgg	1920
58 aagagcgcag	gggtgtaatt	gaatttcacg	tggttggcaa	tccctcaac	cagaaaccaa	1980
59 aacaagaagat	cctgatgtgg	ctgggtggcc	tacagaacgt	tttctccac	cagctgcccc	2040
60 gaatgccaaa	agaatacatc	acacggctcg	tctttgaccc	gaaacacaaa	acccttgctt	2100
61 taattaaaga	tggccgtgtt	attgggtggta	tctgttccg	tatgttcca	tctcaaggat	2160
62 tcacagagat	tgtcttctgt	gctgtaacct	caaatgagca	agtcaagggc	tatgaacac	2220
63 acctgatgaa	tcatttggaa	gaatatcaca	taaagcatga	cacccgtgaac	ttccctcacat	2280
64 atgcagatga	atatgcattt	ggatacttta	agaaacaggg	tttctccaaa	gaaattaaaa	2340
65 tacctaaaac	caaatatgtt	ggctatatac	aggattatga	aggagccact	ttaatggat	2400
66 gtgagctaaa	tccacggatc	ccgtacacag	aattttctgt	catcattaaa	aagcagaagg	2460
67 agataattaa	aaaactgatt	gaaagaaaaac	aggcacaat	tcgaaaagg	tacccgtggac	2520
68 tttcatgttt	taaagatgga	gttcgacaga	ttcctataga	aacgattct	ggaatttagag	2580
69 agacaggctg	gaaaccgggt	ggaaaagaga	aaagtaaaga	gcccagagac	cctgaccagc	2640
70 tttacagcac	gctcaagagc	atccctccagc	aggtgaagag	ccatcaaagc	gttggccct	2700
71 tcatggAAC	tgtgaagaga	acagaagctc	caggatatta	tgaagttata	aggttccca	2760
72 tggatctgaa	aaccatgagt	gaacgcctca	agaataggt	ctacgtgtct	aagaaattat	2820
73 tcatggcaga	cttacagcga	gtctttacca	attgcaaaga	gtacaacgcc	gctgagatg	2880
74 aatactacaa	atgtgcaat	atccctggaga	aattttctt	cagtaaaaatt	aaggaagctg	2940
75 gattaattga	caagtgattt	tttttccccc	tctgcttctt	agaaaactcac	caagcagtgt	3000
76 gcctaaagca	aggt					3014

79 <210> SEQ ID NO: 2

80 <211> LENGTH: 832

81 <212> TYPE: PRT

82 <213> ORGANISM: Homo sapiens

84 <400> SEQUENCE: 2

85 Met Ser Glu Ala Gly	Gly Ala Gly Pro	Gly Gly Cys	Gly Ala Gly Ala	
86 1	5	10	15	
88 Gly Ala Gly Ala Gly	Pro Gly Ala Leu Pro	Pro Pro Gln Pro	Ala Ala Leu	
89	20	25	30	
91 Pro Pro Ala Pro Pro	Gln Gly Ser Pro	Cys Ala Ala Ala	Gly Gly	
92	35	40	45	
94 Ser Gly Ala Cys Gly	Pro Ala Thr Ala Val	Ala Ala Gly Thr	Ala	
95	50	55	60	
97 Glu Gly Pro Gly Gly	Gly Ser Ala Arg Ile	Ala Val Lys Lys	Ala	
98 65	70	75	80	
100 Gln Leu Arg Ser Ala	Pro Arg Ala Lys Lys	Leu Glu Lys Leu	Gly Val	
101	85	90	95	
103 Tyr Ser Ala Cys Lys	Ala Glu Glu Ser	Cys Lys Cys Asn	Gly Trp Lys	
104	100	105	110	
106 Asn Pro Asn Pro Ser	Pro Thr Pro Pro Arg	Ala Asp Leu Gln	Gln Ile	
107	115	120	125	
109 Ile Val Ser Leu Thr	Glu Ser Cys Arg Ser	Cys Ser His	Ala Leu Ala	
110	130	135	140	
112 Ala His Val Ser His	Leu Glu Asn Val	Ser Glu Glu	Met Asn Arg	
113 145	150	155	160	
115 Leu Leu Gly Ile Val	Leu Asp Val Glu	Tyr Leu Phe	Thr Cys Val His	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,553D

DATE: 07/26/2004

TIME: 14:06:30

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
 Output Set: N:\CRF4\07262004\I784553D.raw

116	165	170	175
118	Lys Glu Glu Asp Ala Asp Thr Lys Gln Val Tyr Phe Tyr Leu Phe Lys		
119	180	185	190
121	Leu Leu Arg Lys Ser Ile Leu Gln Arg Gly Lys Pro Val Val Glu Gly		
122	195	200	205
124	Ser Leu Glu Lys Lys Pro Pro Phe Glu Lys Pro Ser Ile Glu Gln Gly		
125	210	215	220
127	Val Asn Asn Phe Val Gln Tyr Lys Phe Ser His Leu Pro Ala Lys Glu		
128	225	230	235
130	Arg Gln Thr Ile Val Glu Leu Ala Lys Met Phe Leu Asn Arg Ile Asn		
131	245	250	255
133	Tyr Trp His Leu Glu Ala Pro Ser Gln Arg Arg Leu Arg Ser Pro Asn		
134	260	265	270
136	Asp Asp Ile Ser Gly Tyr Lys Glu Asn Tyr Thr Arg Trp Leu Cys Tyr		
137	275	280	285
139	Cys Asn Val Pro Gln Phe Cys Asp Ser Leu Pro Arg Tyr Glu Thr Thr		
140	290	295	300
142	Gln Val Phe Gly Arg Thr Leu Leu Arg Ser Val Phe Thr Val Met Arg		
143	305	310	315
145	Arg Gln Leu Leu Glu Gln Ala Arg Gln Glu Lys Asp Lys Leu Pro Leu		
146	325	330	335
148	Glu Lys Arg Thr Leu Ile Leu Thr His Phe Pro Lys Phe Leu Ser Met		
149	340	345	350
151	Leu Glu Glu Glu Val Tyr Ser Gln Asn Ser Pro Ile Trp Asp Gln Asp		
152	355	360	365
154	Phe Leu Ser Ala Ser Ser Arg Thr Ser Gln Leu Gly Ile Gln Thr Val		
155	370	375	380
157	Ile Asn Pro Pro Pro Val Ala Gly Thr Ile Ser Tyr Asn Ser Thr Ser		
158	385	390	395
160	Ser Ser Leu Glu Gln Pro Asn Ala Gly Ser Ser Ser Pro Ala Cys Lys		
161	405	410	415
163	Ala Ser Ser Gly Leu Glu Ala Asn Pro Gly Glu Lys Arg Lys Met Thr		
164	420	425	430
166	Asp Ser His Val Leu Glu Glu Ala Lys Lys Pro Arg Val Met Gly Asp		
167	435	440	445
169	Ile Pro Met Glu Leu Ile Asn Glu Val Met Ser Thr Ile Thr Asp Pro		
170	450	455	460
172	Ala Ala Met Leu Gly Pro Glu Thr Asn Phe Leu Ser Ala His Ser Ala		
173	465	470	475
175	Arg Asp Glu Ala Ala Arg Leu Glu Glu Arg Arg Gly Val Ile Glu Phe		
176	485	490	495
178	His Val Val Gly Asn Ser Leu Asn Gln Lys Pro Asn Lys Lys Ile Leu		
179	500	505	510
181	Met Trp Leu Val Gly Leu Gln Asn Val Phe Ser His Gln Leu Pro Arg		
182	515	520	525
184	Met Pro Lys Glu Tyr Ile Thr Arg Leu Val Phe Asp Pro Lys His Lys		
185	530	535	540
187	Thr Leu Ala Leu Ile Lys Asp Gly Arg Val Ile Gly Gly Ile Cys Phe		
188	545	550	555
			560

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/784,553D

DATE: 07/26/2004
TIME: 14:06:30

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
Output Set: N:\CRF4\07262004\I784553D.raw

190 Arg Met Phe Pro Ser Gln Gly Phe Thr Glu Ile Val Phe Cys Ala Val
191 565 570 575
193 Thr Ser Asn Glu Gln Val Lys Gly Tyr Gly Thr His Leu Met Asn His
194 580 585 590
196 Leu Lys Glu Tyr His Ile Lys His Asp Ile Leu Asn Phe Leu Thr Tyr
197 595 600 605
199 Ala Asp Glu Tyr Ala Ile Gly Tyr Phe Lys Lys Gln Gly Phe Ser Lys
200 610 615 620
202 Glu Ile Lys Ile Pro Lys Thr Lys Tyr Val Gly Tyr Ile Lys Asp Tyr
203 625 630 635 640
205 Glu Gly Ala Thr Leu Met Gly Cys Glu Leu Asn Pro Arg Ile Pro Tyr
206 645 650 655
208 Thr Glu Phe Ser Val Ile Ile Lys Lys Gln Lys Glu Ile Ile Lys Lys
209 660 665 670
211 Leu Ile Glu Arg Lys Gln Ala Gln Ile Arg Lys Val Tyr Pro Gly Leu
212 675 680 685
214 Ser Cys Phe Lys Asp Gly Val Arg Gln Ile Pro Ile Glu Ser Ile Pro
215 690 695 700
217 Gly Ile Arg Glu Thr Gly Trp Lys Pro Ser Gly Lys Glu Lys Ser Lys
218 705 710 715 720
220 Glu Pro Arg Asp Pro Asp Gln Leu Tyr Ser Thr Leu Lys Ser Ile Leu
221 725 730 735
223 Gln Gln Val Lys Ser His Gln Ser Ala Trp Pro Phe Met Glu Pro Val
224 740 745 750
226 Lys Arg Thr Glu Ala Pro Gly Tyr Tyr Glu Val Ile Arg Phe Pro Met
227 755 760 765
229 Asp Leu Lys Thr Met Ser Glu Arg Leu Lys Asn Arg Tyr Tyr Val Ser
230 770 775 780
232 Lys Lys Leu Phe Met Ala Asp Leu Gln Arg Val Phe Thr Asn Cys Lys
233 785 790 795 800
235 Glu Tyr Asn Ala Ala Glu Ser Glu Tyr Tyr Lys Cys Ala Asn Ile Leu
236 805 810 815
238 Glu Lys Phe Phe Ser Lys Ile Lys Glu Ala Gly Leu Ile Asp Lys
239 820 825 830
242 <210> SEQ ID NO: 3
243 <211> LENGTH: 25
244 <212> TYPE: PRT
245 <213> ORGANISM: Artificial Sequence
247 <220> FEATURE:
248 <223> OTHER INFORMATION: synthetic bromodomain peptide
250 <220> FEATURE:
251 <221> NAME/KEY: MOD_RES
252 <222> LOCATION: (2)..(4)
253 <223> OTHER INFORMATION: Any amino acid; this range may encompass 2-3 residues
255 <220> FEATURE:
256 <221> NAME/KEY: MOD_RES
257 <222> LOCATION: (6)..(13)
258 <223> OTHER INFORMATION: Any amino acid; this range may encompass 5-8 residues
260 <220> FEATURE:

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/784,553D

DATE: 07/26/2004
TIME: 14:06:30

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
Output Set: N:\CRF4\07262004\I784553D.raw

261 <221> NAME/KEY: MOD_RES
 262 <222> LOCATION: (14) /
 263 <223> OTHER INFORMATION: Pro, Lys or His
 265 <220> FEATURE:
 266 <221> NAME/KEY: MOD_RES
 267 <222> LOCATION: (15) /
 268 <223> OTHER INFORMATION: Any amino acid
 270 <220> FEATURE:
 271 <221> NAME/KEY: MOD_RES
 272 <222> LOCATION: (17) /
 273 <223> OTHER INFORMATION: Tyr, Phe or His
 275 <220> FEATURE:
 276 <221> NAME/KEY: MOD_RES
 277 <222> LOCATION: (18)..(22) /
 278 <223> OTHER INFORMATION: Any amino acid
 280 <220> FEATURE:
 281 <221> NAME/KEY: MOD_RES
 282 <222> LOCATION: (24) /
 283 <223> OTHER INFORMATION: Met, Ile or Val
 285 <400> SEQUENCE: 3
W--> 286 Phe Xaa Xaa Xaa Pro Xaa Tyr
 287 1 5 10 15
 289 Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Asp
 290 20 25
 292 <210> SEQ ID NO: 4
 293 <211> LENGTH: 12
 294 <212> TYPE: PRT
 295 <213> ORGANISM: Artificial Sequence
 297 <220> FEATURE:
 298 <223> OTHER INFORMATION: synthetic bromodomain peptide
 300 <220> FEATURE:
 301 <221> NAME/KEY: MOD_RES
 302 <222> LOCATION: (6)..(6) /
 303 <223> OTHER INFORMATION: acetyl lysine
 305 <400> SEQUENCE: 4
W--> 306 Ile Ser Tyr Gly Arg Xaa Lys Arg Arg Gln Arg Arg
 307 1 5 10
 309 <210> SEQ ID NO: 5
 310 <211> LENGTH: 14
 311 <212> TYPE: PRT
 312 <213> ORGANISM: Artificial Sequence
 314 <220> FEATURE:
 315 <223> OTHER INFORMATION: synthetic bromodomain peptide
 318 <220> FEATURE:
 319 <221> NAME/KEY: MOD_RES
 320 <222> LOCATION: (8)..(8) /
 321 <223> OTHER INFORMATION: acetyl lysine
 323 <400> SEQUENCE: 5
W--> 324 Ala Arg Lys Ser Thr Gly Gly Xaa Ala Pro Arg Lys Gln Leu

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/26/2004
PATENT APPLICATION: US/09/784,553D TIME: 14:06:31

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
Output Set: N:\CRF4\07262004\I784553D.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos. 2,3,4,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24
Seq#:4; Xaa Pos. 6
Seq#:5; Xaa Pos. 8
Seq#:6; Xaa Pos. 8
Seq#:43; Xaa Pos. 1,2,4,5,6,8,9,10,11,12,13,14,15,16,17,19,20,21,22,23,24
Seq#:43; Xaa Pos. 26
Seq#:46; Xaa Pos. 5,6,7
Seq#:48; Xaa Pos. 2,3,4,5,7,8,9,10,12,13,14,15,17,18,19,21
Seq#:50; Xaa Pos. 5
Seq#:51; Xaa Pos. 5
Seq#:52; Xaa Pos. 5
Seq#:53; Xaa Pos. 5
Seq#:54; Xaa Pos. 5
Seq#:55; Xaa Pos. 5
Seq#:56; Xaa Pos. 5
Seq#:57; Xaa Pos. 6
Seq#:58; Xaa Pos. 7
Seq#:59; Xaa Pos. 16
Seq#:62; Xaa Pos. 8
Seq#:63; Xaa Pos. 7

VERIFICATION SUMMARY DATE: 07/26/2004
PATENT APPLICATION: US/09/784,553D TIME: 14:06:31

Input Set : A:\2459-1-003CIP Substitute SEQ LIST.txt
Output Set: N:\CRF4\07262004\I784553D.raw

L:286 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
M:341 Repeated in SeqNo=3
L:306 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:324 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:1440 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0
M:341 Repeated in SeqNo=43
L:1507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0
L:1558 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0
M:341 Repeated in SeqNo=48
L:1601 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0
L:1618 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0
L:1635 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0
L:1652 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53 after pos.:0
L:1669 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0
L:1686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0
L:1703 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0
L:1720 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:0
L:1737 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:0
L:1755 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:0
L:1801 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62 after pos.:0
L:1819 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63 after pos.:0